Volume 317, number 1,2, 96–100 FEBS 12079 February 1993

Structure of the 16 S ribosomal RNA of the thermophilic cyanobacterium *Chlorogloeopsis* HTF ('*Mastigocladus laminosus* HTF') strain PCC7518, and phylogenetic analysis

Annick Wilmotte, Gert Van der Auwera and Rupert De Wachter

Departement Biochemie, Universiteit Antwerpen (UIA), Universiteitsplein 1, B-2610 Antwerpen, Belgium

Received 10 November 1992

The thermophilic cyanobacterial strain, PCC7518, originally identified as '*Mastigocladus laminosus* HTF' does not show branchings or heterocysts. The absence of branchings supports the later assignment to the genus *Chlorogloeopsis*. The absence of heterocysts may be the result of a mutation because heterocysts were observed in the original isolate. Alternatively, contamination may have happened. To solve this problem, the 16 S rRNA sequence was determined and used to infer a secondary structure model and build distance trees. The trees showed that strain PCC7518 belongs to the cluster of heterocystous species and has most probably lost the ability to produce heterocysts by mutation. It is only distantly related to *Chlorogloeopsis fritschii* PCC6718.

16 S ribosomal RNA; Cyanobacterial evolution; Blue-green algae; Heterocyst; *Mastigocladus laminosus* HTF; *Chlorogloeopsis*

1. INTRODUCTION

'*Mastigocladus laminosus* HTF' clone I-15 was isolated by Castenholz [1,2] in Iceland, from a small pond the temperature of which varied between 62 and 73°C and had a pH of 9. It was later sent to the Pasteur Culture Collection (Paris, France) where it was assigned the strain number, PCC7518. The cells of strain PCC7518 form irregular aggregates and short filaments without branchings (Rippka, personal communication). Originally, clone I-15 was identified by Castenholz [1,2] as a form of *Mastigocladus laminosus*, following the taxonomic systems of Frémy [3] and Schwabe [4] who assumed the existence of several forms and subforms in this species. Later, Castenholz [5] assigned *Mastigocladus laminosus* HTF to the genus *Chlorogloeopsis* on the basis of their morphological similarity. Indeed, the latter genus is characterized by the presence of cell aggregates and short filaments without branchings, the hormogonia. A peculiarity of strain PCC7518 is that it never forms heterocysts, although these had been observed in clone I-15 [1,2]. Because of the uncertain taxonomic affiliation, this strain was not recorded in a taxonomic system [6] of the cyanobacteria based on the strains of the Pasteur Culture Collection. It was thus of interest to assess the genotypic relationships of this enigmatic strain to other cyanobacteria by means of 16 S rRNA sequence analysis.

Correspondence address: R. De Wachter, Departement Biochemie, Universiteit Antwerpen (UIA), Universiteitsplein 1, B-2610 Antwerp, Belgium. Fax: (32) (3) 820 2248.

2. MATERIALS AND METHODS

2.1. *Gene amplification by the polymerase chain reaction (PCR)*

Strain PCC7518 was obtained from the Pasteur Culture Collection, Pasteur Institute, Paris. About 0.5 μl of a green liquid at the surface of the agar slant was added directly to 100 μl of PCR-mix containing 2.8 ng/μl of primers 1 and 18 (Table 1), 200 μM of each DNTP, 2.5 U of *Taq* DNA polymerase, 10 mM Tris-Hcl, 50 mM KCl, 1.5 mM $MgCl_2$, 0.01% gelatin (w/v); pH 8.3. After an initial denaturation of 1 min at 95°C, 30 cycles of amplification were started: 1 min at 95°C, 1 min at 55°C and 4 min at 72°C. A final incubation of 7 min at 72°C was carried out.

2.2. *Cloning and sequencing*

The PCR product, after purification by electrophoresis on agarose gels, was cloned in a vector, pSK^+ (Stratagene) to which a ddT-tail had been added at the 3'-terminus by terminal transferase [10]. After transformation into competent *E. coli* DH5α cells, colony hybridization was carried out [11]. The probe was the PCR product labelled with ^{32}P using a nick-translation kit (Amersham). Sequence determination was carried out with the Sequenase 2.0 sequencing kit (USB) according to the manufacturer's instructions. The two M13 primer sites on the vector and eighteen primers annealing to evolutionarily conserved areas, listed in Table I, were used to sequence both strands of the gene as well as a part of the internal transcribed spacer (data not shown).

2.3. *Alignment and distance tree construction*

The 16 S rRNA sequence was aligned with all other known complete or nearly complete bacterial sequences, according to the primary and secondary structure features [12]. A matrix of distances, corrected for multiple mutations, was calculated as previously described [13] and a tree was constructed by means of the Neighbor-Joining method [14] using the software package Treecon [15]. The statistical support for the branches was assessed by means of a bootstrap analysis [16] involving the construction of 100 new trees from re-sampled data.

 Published by Elsevier Science Publishers B.V

Table I

Oligodeoxyribonucleotides used as primers for PCR and sequence determination

Number	Sequence[a]	Strand[b]	Positions in the corresponding genes of *Synechococcus* sp. PCC6301 [7–9]
1	AGAGTTTGATCCTGGCTCAG	R	8–27 (16 S)
2	TCGACTTGCATGTRTTA	C	65–49 (16 S)
3	TACGGRAGGCAGCAG	R	310–324 (16 S)
4	AGGCAGCAGTGGGGAAT	R	316–332 (16 S)
5	CTGCTGCCTYCCGTA	C	324–310 (16 S)
6	GCCAGCAGCCGCGGTAA	R	460–476 (16 S)
7	ATTACCGCGGCTGCTGGC	C	477–460 (16 S)
8	RGGATTAGATACCCC	R	727–741 (16 S)
9	GGGGTATCTAATCCC	C	741–727 (16 S)
10	GAATTGACGGGGRCCC	R	861–876 (16 S)
11	CCGTCAATTYYTTTRAGTTT	C	870–851 (16 S)
12	GYAACGAGCGCAACCC	R	1043–1058 (16 S)
13	GGGTTGCGCTCGTTRC	C	1058–1043 (16 S)
14	TGTACACACCGCCCGTC	R	1334–1350 (16 S)
15	GACGGGCGGTGTGTRC	C	1350–1335 (16 S)
16	AAGGAGGTGATCCAGCCGCA	C	1486–1467 (16 S)
17	ATTAGCTCAGGTGGTTAG	R	6–23 ($tRNA^{Tle}$)
18	TTTGCGGCCGCTCTGTGTGCCTAGGTATCC	C	26–45 (23 S)

[a]The simultaneous presence of two bases at a single position is indicated by R (A + G) and Y (C + T).
[b]Oligonucleotides with a sequence corresponding to the RNA-like strand are indicated by R; oligonucleotides complementary to the RNA-like strand are indicated by C.

3. RESULTS AND DISCUSSION

3.1. *Primary and secondary structure of the 16 S rRNA*

The PCR product has a length of about 2.1 kb. Ten recombinant plasmids were obtained and pooled to determine a sequence of 1,461 nucleotides, which is the complete 16 S rDNA sequence, except for an estimated 27 nucleotides extending from the 5′-terminus to the 3′ end of primer 1 (Table I). The accession number of the gene in the EMBL data bank is X68780. Its secondary structure is given in Fig. 1. The most notable change with respect to the secondary structure model proposed for *Synechococcus* sp. PCC6301 ('*Anacystis nidulans*') [17] is that the present model takes into account the tertiary interaction involving helices 19, 20 and 21 [18]. In addition, slightly different base pairing schemes are assumed for helices 6, 44 and 46.

3.2. *Phylogenetic analysis*

Fig. 2a shows a distance tree constructed from all available complete cyanobacterial sequences [7,19]. However, partial sequences have been determined for many more cyanobacterial strains. For construction of the tree of Fig. 2b, all partial sequences of heterocystous strains [22,23] were selected, and the strains represented in Fig. 2a were included as well. Both trees show clearly that strain PCC7518 belongs to the cluster of heterocystous strains. This grouping is supported by a bootstrap value of 100%. The genomic DNA base composition of 44.8% (Herdman, personal communication) is also consistent with the assignment of strain PCC7518 to the cluster of heterocystous species [24]. Accordingly, it is probable that strain PCC7518 has lost the ability to produce heterocysts by mutation. Cases of such spontaneous losses have been observed in *Nostoc* or *Anabaena* strains which had been cultivated for a long time in the presence of nitrates (Rippka, personal communication). Thus, it would be interesting to study the presence and arrangement of *nif* genes and compare them with the data obtained for *Mastigocladus laminosus* UTEX1931 [25].

In the last edition of Bergey's Manual of Systematic Bacteriology, *Mastigocladus laminosus* HTF is assigned to the genus *Chlorogloeopsis*, on the basis of its morphology [5]. However, the 16 S rRNA similarity of strain PCC7518 with *Chlorogloeopsis fritschii* PCC6718 is only 91.1% for the 992 nucleotides known for the two sequences. The sequence similarity is even slightly lower with *Fischerella* PCC7414, namely 89.6% for the 1,002 nucleotides available. These values, and the topology of divergence observed in Fig. 2b, demonstrate that strain PCC7518 should be assigned to a genus different from those to which *Chlorogloeopsis fritschii* PCC6718 or *Fischerella* PCC7414 belong. For a more precise taxonomic ranking of the three strains, DNA–DNA hybridization data would be useful [26]. In Fig. 2b, the lineage to which strain PCC7518 belongs shares a short common stem with *Fischerella* PCC7414 and *Chlorogloeopsis fritschii* PCC6718 but diverges rather early. The cluster of these three strains corresponds to the Section V proposed by Rippka et al. [6] and to the order, Stigonematales [5]. However, the bootstrap percentage sup-

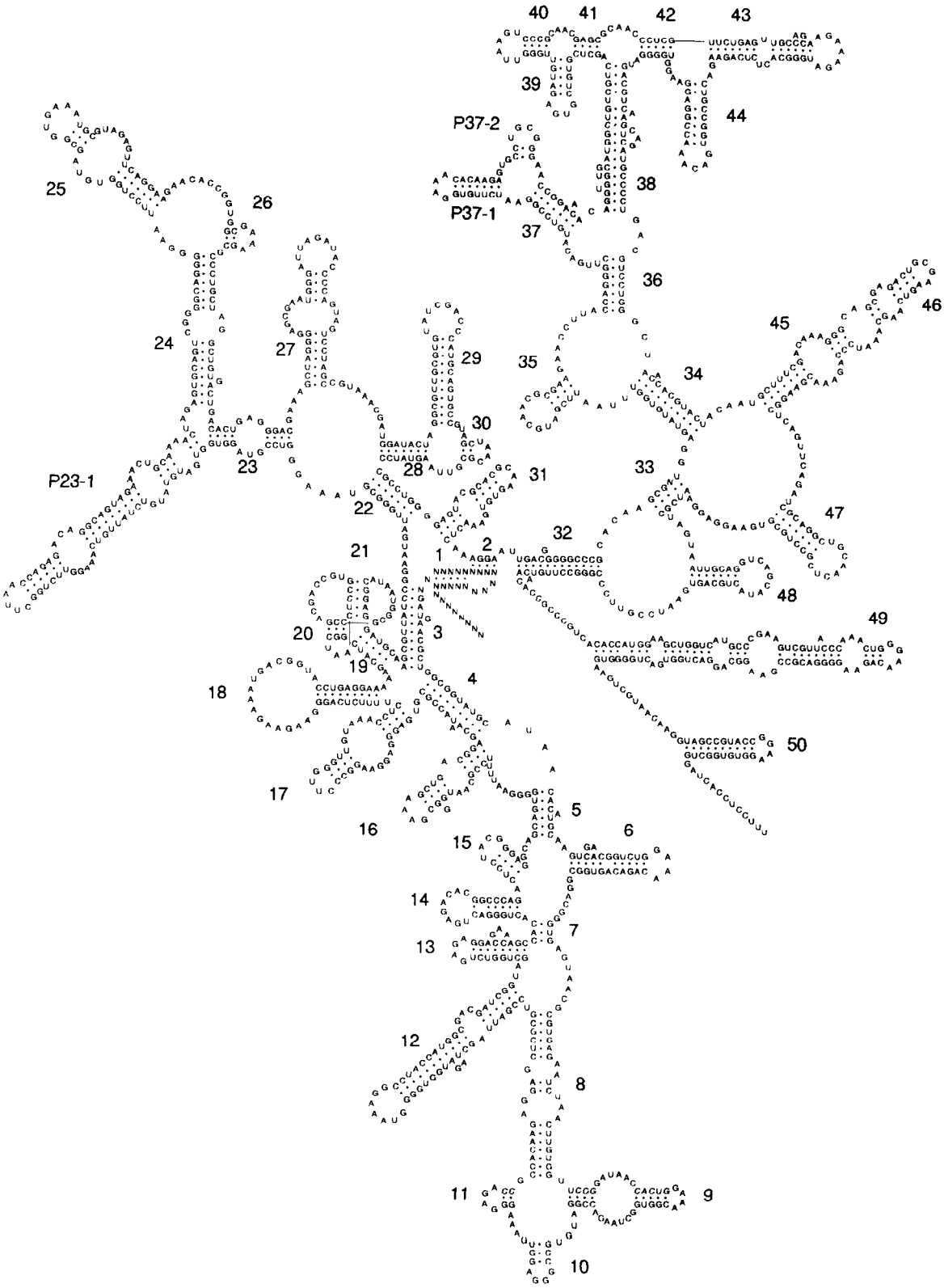

Fig. 1. Secondary structure model for 16 S rRNA of strain PCC7518. The number of undetermined nucleotides at the 5'-terminus, indicated as 'N', was estimated by comparison with the complete known sequence of *Synechococcus* PCC6301 [7]. The tertiary interaction involving helices 19, 20 and 21 results in a change in the helix numbering system with respect to the secondary structure model for prokaryotic small subunit rRNA of De Rijk et al. [12].

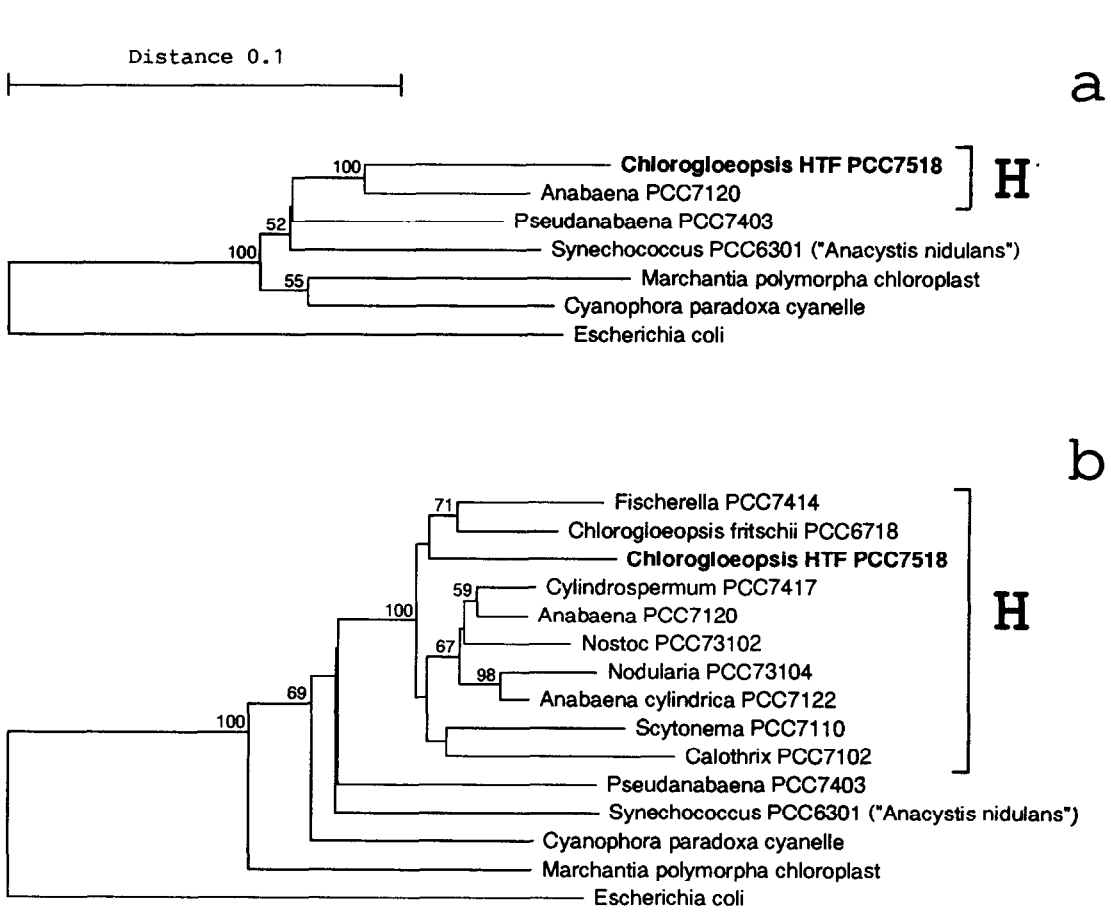

Fig. 2. Distance trees based on 16 S rRNA sequence analysis. Bootstrap percentages higher than 50% are indicated for each node. Clusters consisting of heterocystous strains are indicated as 'H'. (a) Tree based on the complete 16 S rRNA sequences of strain PCC7518, the other complete cyanobacterial sequences published to date [7,19], the sequence of *Pseudanabaena* PCC7403 (unpublished), *Marchantia polymorpha* plastid [20], *Cyanophora paradoxa* cyanelle (D. Bryant, personal communication) and *E. coli* [21] which was used to root the tree. (b) Tree based on all the partial 16 S rRNA sequences from heterocystous cyanobacteria published to date [22,23] and corresponding sequence sections of additional strains included in Fig. 2a. The alignment included 763 positions.

porting this cluster is lower than 50%. This indicates a certain instability within the heterocystous cluster and some caution is necessary while interpreting the tree topology. The future examination of 16 S rRNA from other closely related organisms may help to stabilize the position of the studied strain in the distance tree.

Acknowledgements: The authors thank Raymond De Baere for technical advice, Jean-Marc Neefs and Peter De Rijk for the alignment, Yves Van de Peer for help with one figure, Rosi Rippka (Pasteur Institute, Paris) for the idea of this research and useful information, Richard Castenholz (Oregon University) for valuable information and Donald Bryant (Pennsylvania University) for an unpublished sequence. This research was supported by the Fund for Medical Scientific Research (FGWO).

REFERENCES

[1] Castenholz, R.W. (1969) J. Phycol. 5, 360–368.
[2] Castenholz, R.W. (1978) Mitt. Int. Verein. Limnol. 21, 296–315.
[3] Frémy, P. (1936) Ann. Protist. 5, 175–200.
[4] Schwabe, G.H. (1960) Schweiz. Z. Hydrol. 22, 757–792.
[5] Castenholz, R.W. (1989) in: Bergey's Manual of Systematic Bacteriology, vol. 3 (Staley, J.T., Bryant, M.P., Pfennig, N. and Holt, J.G. eds.) pp. 1794–1799, Williams and Wilkins, Baltimore, USA.
[6] Rippka, R., Deruelles, J., Waterbury, J.B., Herdman, M. and Stanier, R.Y. (1979) J. Gen. Microbiol. 111, 1–61.
[7] Tomioka, N. and Sugiura, M. (1983) Mol. Gen. Genet. 191, 46–50.
[8] Kumano, M., Tomioka, N. and Sugiura, M. (1983) Gene 24, 219–225.
[9] Tomioka, N. and Sugiura, M. (1984) Mol. Gen. Genet. 193, 427–430.
[10] Holton, T.A. and Graham, M.W. (1991) Nucleic Acids Res. 19, 1156.
[11] Dale, J.W. and Greenaway, P.J. (1984) in: Methods in Molecular Biology, vol. 2 (Walker, J.M. ed.) pp. 277–283, Humana Press, Clifton, NJ, USA.
[12] De Rijk, P., Neefs, J.M., Van de Peer, Y. and De Wachter, R. (1992) Nucleic Acids Res. 20, 2075–2089.
[13] Van de Peer, Y., Neefs, J.-M. and De Wachter, R. (1990) J. Mol. Evol. 30, 463–476.
[14] Saitou, N. and Nei, M. (1987) Mol. Biol. Evol. 4, 406–425.
[15] Van de Peer, Y. and De Wachter, R. (in press) Comput. Appl. Biosc.
[16] Felsenstein, J. (1985) Evolution 39, 783–791.

[17] Gutell, R.R., Weiser, B., Woese, C.R. and Noller, H.F. (1985) Prog. Nucleic Acids Res. Mol. Biol. 32, 155–216.
[18] Powers, T. and Noller, H.F. (1990) Proc. Natl. Acad. Sci USA 87, 1042–1046.
[19] Ligon, P.J.B., Meyer, K.G., Martin, J.A. and Curtis, S.E. (1991) Nucleic Acids Res. 19, 4453.
[20] Ohyama, K., Fukuzawam H., Kohchi, T., Shirai, H., Sano, T., Sano, S., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S., Inokuchi, H. and Ozeki, H. (1986) Nature 322, 572–574.
[21] Brosius, J., Dull, T., Sleeter, D.D. and Noller, H.F. (1981) J. Mol. Biol. 148, 107–127.
[22] Giovannoni, S.J., Turner, S., Olsen, G.J., Barns, S., Lane, D.J. and Pace, N.R. (1988) J. Bacteriol. 170, 3584–3592.
[23] Wilmotte, A., Turner, S., Van de Peer, Y. and Pace, N.R. (1992) J. Phycol. 28, 828–838.
[24] Herdmann, M., Janvier, M., Waterbury, J.B., Rippka, R., Stanier, R.Y. and Mandel, M. (1979) J. Gen. Microbiol. 111, 63–71.
[25] Singh, R.K. and Stevens, S.E. (1982) FEMS Microbiol. Lett. 94, 227–234.
[26] Wayne, L.G., Brenner, D.J., Colwell, R.R., Grimont, P.A.D., Kandler, O., Krichevsky, M.I., Moore, L.H., Moore, W.E.C., Murray, R.G.E., Stackebrandt, E., Starr, M.P. and Trüper, H.G. (1987) Int. J. System. Bacteriol. 37, 463–464.